Qn III .[16] **[Social Networks and Power law Phenomena]**

[2] Would the existence of pages like Google and Yahoo!, that have very high in-degree, be **more probable** or **less probable** if the web graph was a uniform-random (also called "exponential") network (rather than a scale-free network)?

[3] Consider a uniform random (aka "exponential") network with 512 nodes. (i) What is the maximum possible diameter of this network? (ii) If I now tell you that this network was generated by each node making a connection with another node with (a uniform) probability of 0.006. What is the expected diameter of the resulting network?

[3] The military is building a new internet that is meant for just the military machines. They are particularly interested in ensuring that the network is resilient against denial of service attacks. Being military, they decided to provide detailed instructions for individual units as to how they should connect themselves to other units. Here are two alternative instructions they are considering. Which one would you advise them to use and why?

a. Each unit should randomly pick k other units and establish network connections to them
b. Each unit should randomly pick another unit, and copy some of the connections that that unit has established.

[2] Consider a corpus of documents. Suppose the word frequency in the documents follows a Zipfian distribution. If the 100^{th} most popular word occurs 1000 times, what is the approximate frequency of the 50^{th} most popular word?

[2] Suppose the corpus above is growing—i.e., we are adding more and more documents to it. We know, by Heap's law, that the lexicon (vocabulary used in the documents) is expected grow at a rate that is square root of the size of the corpus (measured in number of words). Now, Heap's law assumes that the documents are generated by sampling words using a *Zipfian* distribution. How will the lexicon's rate increase change if the documents in the corpus, instead of being generated according to a Zipfian distribution, are generated by *uniform distribution* (You only need to compare to the rate of change w.r.t. what is expected by Heap's law). **Explain**.